

DOCUMENT RESUME

ED 045 730

TM 000 317

AUTHOR Gregory, Thomas B.
TITLE A Computer Program for Calculation of Scott's Coefficient of Observer Reliability.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO R-Ser-21
PUB DATE Sep 69
GRANT USOE-6-10-108
NOTE 10p.

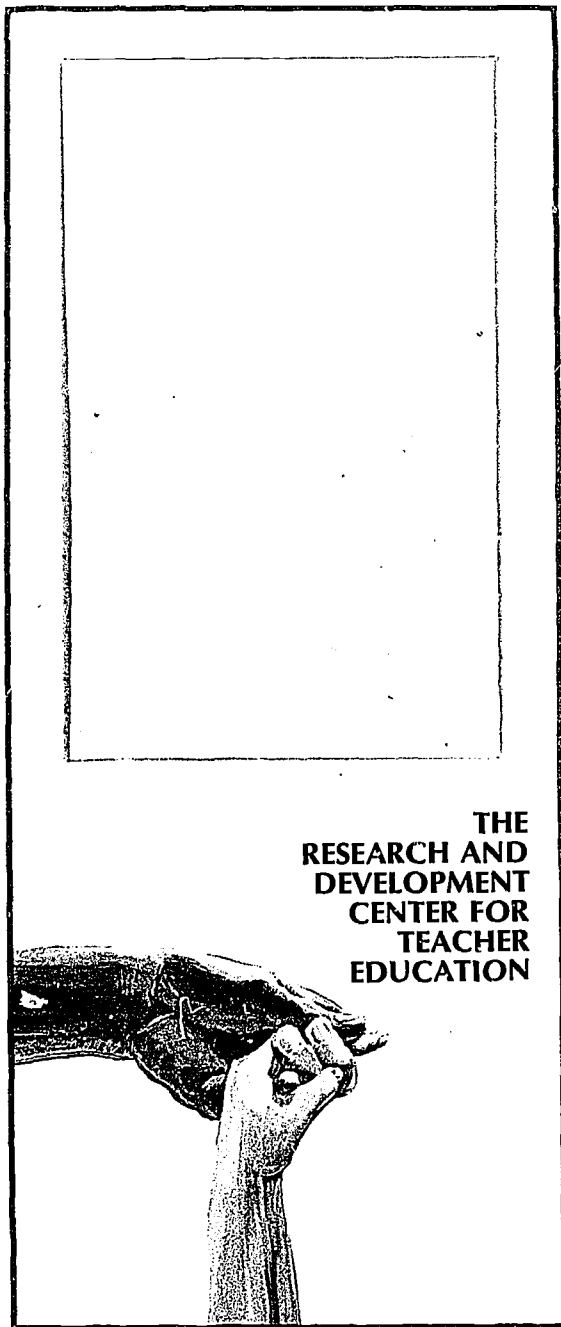
EDRS PRICE EDRS Price MF-\$0.25 HC-\$0.60
DESCRIPTORS *Classroom Observation Techniques, *Computer Programs, Educational Research, *Reliability, *Research Methodology, *Statistical Analysis
IDENTIFIERS FORTRAN, Interrater Reliability, Observer Reliability

ABSTRACT

Scott's coefficient is an improved method for estimating observer reliability using any system which assigns events to mutually exclusive categories. The formula is the ratio of the actual difference between obtained and chance agreement to the maximum difference between obtained and chance agreement. A FORTRAN program for calculating this reliability coefficient is presented. The program can be used with from 2 to 20 observers and can include any number of observations as long as the product of the observations times the observers does not exceed 200. The observation system cannot exceed 30 categories. There are four outputs of the program including the average reliability coefficient of all possible pairs of observers, the average difference between all possible pairs of observers for each category, the percentage of each observer's total assignments found in each category, and each observer's average coefficient of reliability with all other observers. The data deck arrangement and program listing are included. (PR)

ED0 45730

TM 000 317



THE UNIVERSITY OF TEXAS AT
AUSTIN / AUSTIN PUBLIC
SCHOOLS / TEXAS EDU-
CATION AGENCY

ED0 45730

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE

OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED
EXACTLY AS RECEIVED FROM THE PERSON OR
ORGANIZATION ORIGINATING IT. POINTS OF
VIEW OR OPINIONS STATED DO NOT NECES-
SARILY REPRESENT OFFICIAL OFFICE OF EDU-
CATION POSITION OR POLICY.

A COMPUTER PROGRAM
FOR CALCULATION OF
SCOTT'S COEFFICIENT
OF OBSERVER RELIABILITY

Thomas B. Gregory

Report Series No. 21

September, 1969

The Research and Development Center for Teacher Education
The University of Texas at Austin

This research was conducted with partial support of USOE Grant No. 6-10-108. The author wishes to acknowledge the assistance of Dr. Donald J. Veldman, Dr. Marcella Kysilka and Dr. O.L. Davis, Jr., in the preparation of this program.

A COMPUTER PROGRAM FOR CALCULATION OF SCOTT'S
COEFFICIENT OF OBSERVER RELIABILITY

By

Thomas B. Gregory¹
Indiana University

The past decade has witnessed the development of numerous observation systems that attempt to objectively record the ongoing verbal and nonverbal events occurring in the classroom. Flander's Interaction Analysis (1960) and the several revisions of the OScAR (e.g., Medley and Mitzel, 1958; Medley and others, 1968) are notable examples of this trend. Since such descriptive systems increasingly are being utilized as criterion measures in educational research, the problem of obtaining valid estimates of inter-observer reliability is a crucial one.

Scott (1955) proposed a method of estimating the reliability of observers using any observation system that assigns events to mutually exclusive categories. The method marks an

¹Formerly, Department of Curriculum and Instruction, College of Education, The University of Texas at Austin.

improvement over conventional methods that are based on the percentage of observer agreements. Since such methods require the assumption of equi-probability of occurrence for all categories, they are inappropriate for these observation systems. The formula for Scott's coefficient is:

$$\mathcal{K} = \frac{P_o - P_e}{1 - P_e}$$

where P_o (observed percent of agreement) represents the percentage of judgements on which two observers agree when coding the same data independently; and P_e is the percent of agreement to be expected by chance. \mathcal{K} , then, is the ratio of the actual difference between obtained and chance agreement to the maximum difference between obtained and chance agreement (Scott, 1955, p. 323). Detailed procedures for hand-calculation of \mathcal{K} are found elsewhere (Flanders, 1960, 1967).

General Description of the Computer Program

This Fortran program may be used with from 2 to 20 observers. Any number of individual observations may be included as long as the product of the observation $N \times$ the observer N does not exceed 200 (i.e., the maximum observation N for 20 observers is 10). The observation system cannot exceed 30 categories. Data may be input as decimal fractions or percents of the total assignments by each observer or as raw scores for each category for each observer. Any input format may be specified. Subroutine CCDS (Veldman, 1967, p. 132) allows several separate jobs to be "stacked" if the need arises.

Four outputs of the program are:

1. The average difference between all possible pairs of observers for each category.
2. The percentage of each observer's total assignments found in each category.
3. The average reliability coefficient of all possible pairs of observers.
4. Each observer's average coefficient of reliability with all other observers.

Data Deck Arrangement

Card 1 Title Control Card

Columns:

1-80 Title identifying this particular job. It may be of any length but cannot be left blank.

Card 2 Parameter Control Card

Columns:

1-5 Number of observers (max., 20).
 6-10 Number of observations (max., 200 / observer N).
 11-15 Number of categories in the system (max., 30).
 16-20 A 1 if data is input as percents of total,
 a 2 if data is input as decimal fractions of total, or
 a 3 if data is input as raw scores for each category.
 21-80 Blank.

Card 3

Columns:

1-80 An appropriate format statement (F mode) for reading the data as it is arranged on the tab cards (e.g., (10X, 10F3)).

- Card 4 Beginning of the data deck arranged with all of
observer 1's observations first, then all of ob-
server 2's,...then all of observer N's.
- Last Card Another Title Control Card indicating the beginning
of a new job or a blank card.

Program Listing

REFERENCES

- Flanders, Ned A. Interaction Analysis in the Classroom. Minneapolis: University of Minnesota, College of Education, (Pre-Publication draft), 1960. (Mimeographed).
- _____. "The Problems of Observer Training and Reliability." pp. 158-66 in Amidon, Edmund J. and Hough, John B. Interaction Analysis: Theory, Research, and Application. Reading, Massachusetts: Addison-Wesley Publishing Co., 1967.
- Medley, Donald M. and Mitzel, Harold E. "A Technique for Measuring Classroom Behavior." Journal of Educational Psychology, 49: 86-92; 1958.
- Medley, Donald M., Schluck, Carolyn G. and Ames, Nancy P. Assessing the Learning Environment in the Classroom: A Manual for Users of OScAR 5V. Princeton: Educational Testing Service, 1968. (Mimeographed).
- Scott, William A. "Reliability of Content Analysis: The Case for Nominal Scale Coding." Public Opinion Quarterly, 19: 321-25; Fall, 1955.
- Veldman, Donald J. Fortran Programming for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1967.

PROGRAM SCOTTR (INPUT, OUTPUT)

C

```

ODIMENSION R(30,200), RT(30,20), PI(20,20), S(200), DIF(30),
1AVG(30), ADIF(30), RAVG(20), KF(16)
5 CALL CCDS (KF, NR, NT, NC, NP, N)
  TPI = 0
  DO 15 I = 1,NC
    DO 10 J = 1,NR
10 RT(I,J) = 0
15 ADIF(I) = 0
    DO 20 I = 1,NR
20 RAVG(I) = 0
    NRT = NR * NT
    READ KF, ((R(I,J), I = 1,NC), J = 1,NRT)
    DO 25 I = 1,NC
    DO 25 J = 1,NRT
25 R(I,J) = R(I,J) + 0.00001
    GO TO (55, 45, 30) NP
30 DO 35 I = 1,NC
    DO 35 J = 1,NRT
35 S(J) = SUMF (R, J, NC, 30)
    DO 40 J = 1,NRT
    DO 40 I = 1,NC
40 R(I,J) = R(I,J) / S(J)
45 DO 50 J = 1,NRT
    DO 50 I = 1,NC
50 R(I,J) = R(I,J) * 100
55 NT1 = 1
    NT2 = NT
    DO 70 J = 1,NR
    DO 65 I = 1,NC
    DO 60 K = NT1, NT2
60 RT(I,J) = RT(I,J) + R(I,K)
65 RT(I,J) = RT(I,J) / NT
    NT1 = NT1 + NT
70 NT2 = NT2 + NT
    DO 85 I = 1,NR
    DO 85 J = 1,NR
    IF (I .NE. J) GO TO 75
    PI(I,J) = 1
    GO TO 85
75 DO 80 K = 1,NC
    DIF(K) = ABS (RT(K,I) - RT(K,J))
    ADIF(K) = ADIF(K) + DIF(K)
80 AVG(K) = (((RT(K,I) + RT(K,J)) / 2) ** 2) * .01
    DIFS = SUMF (DIF, 1, NC, 30)
    AVGS = SUMF (AVG, 1, NC, 30)
    PI(I,J) = (100 - DIFS - AVGS) / (100 - AVGS)
85 TPI = TPI + PI(I,J)
DO 90 K = 1,NC
90 ADIF(K) = ADIF(K) / (NR * NR - NR)
    TPI = (TPI - NR) / (NR * NR - NR)
    DO 100 I = 1,NR
    DO 95 J = 1,NR
    IF (I .EQ. J) GO TO 95

```

```

      RAVG(I) = RAVG(I) + PI(I,J)
95  CONTINUE
100 RAVG(I) = RAVG(I) / (NR - 1)
    PRINT 105
1050FORMAT (37H1PERCENT OF TOTAL FOR EACH CATEGORY. // 22H CAT. A.DIF
1.  RATER //)
    PRINT 110, (J, J = 1,NR)
110 FORMAT (13X, 20I5 /)
    DO 120 I = 1,NC
      PRINT 115, I, ADIF(I), (RT(I,J), J = 1,NR)
115 FORMAT (/ 1X, I2, F8.2, 3X, 20F5.1)
120 CONTINUE
    PRINT 125, TPI
125 FORMAT (// 39H0COEFFICIENT FOR ENTIRE RATER GROUP IS , F6.4)
    PRINT 130, NR
1300FORMAT (/ 66H0SCOTTS RELIABILITY COEFFICIENTS FOR ALL POSSIBLE COM
BINATIONS OF , I2, 9H RATERS. /)
    CALL PRTS (PI, NR, NR, 1H , 20)
    PRINT 135
135 FORMAT (// 34H0AVERAGE R WITH ALL OTHER RATERS. //)
    PRINT 140, (I, RAVG(I), I = 1,NR)
140 FORMAT (I3, F10.4 /)
    GO TO 5
  END

```

C
C

SUBROUTINE CCDS (KF, KI, KJ, KK, KL, KM)

C

```

  DIMENSION KF(16), KH(16)
  READ 5, KH
  5 FORMAT (16A5)
  IF (KH(1) .EQ. KH(2)) STOP
  READ 10, KI, KJ, KK, KL, KM, KF
10 FORMAT (5I5 / 16A5)
  PRINT 15, KH, KI, KJ, KK, KL, KM, KF
150FORMAT (1H1, 16A5 // 11H PARAMETERS / 13H COL 1-5 = , I5 /
113H COL 6-10 = , I5 / 13H COL 11-15 = , I5 / 13H COL 16-20 = ,
2I5 / 13H COL 21-25 = , I5 // 15H DATA FORMAT = , 16A5)
  RETURN
  END

```

C
C

SUBROUTINE PRTS (X, N, M, KH, ND)

C

```

  DIMENSION X(ND,M)
  IF (M .GT. 1) GO TO 20
  PRINT 15
  DO 10 I = 1, N, 10
    J = MIN0 (I + 9, N)
    PRINT 5, KH, (X, K = I,J)
  5 FORMAT (/ A7, 10I11)
10 PRINT 15, (X(K,1), K = I,J)
15 FORMAT (10X, 10F11.4)
  RETURN
20 DO 25 K = 1, M, 10

```

```

PRINT 15
L = MINO (K + 9, M)
PRINT 5, KH, (J, J = K,L)
DO 25 I = 1,N
25 PRINT 30, I, (X(I,J), J = K,L)
30 FORMAT (/ 16, 4X, 10F11.4)
RETURN
END

```

CORRIGENDUM

The following function was accidentally omitted from the program listing:

```

C
FUNCTION SUMF (X, KK, NN, ND)
  DIMENSION X(ND,1)
  SUMF = 0
  N = IABS (NN)
  K = IABS (KK)
  IF (NN) 5, 55, 10
  5 IF (KK) 15, 55, 25
  10 IF (KK) 35, 55, 45
  15 DO 20 I = 1,N
  20 SUMF = SUMF + X(K,I) ** 2
  RETURN
  25 DO 30 I = 1,N
  30 SUMF = SUMF + X(I,K) ** 2
  RETURN
  35 DO 40 I = 1,N
  40 SUMF = SUMF + X(K,I)
  RETURN
  45 DO 50 I = 1,N
  50 SUMF = SUMF + X(I,K)
  55 RETURN
END

```